RELIABILITY OF COP-BASED POSTURAL SWAY MEASURES

Dingding Lin, Hyang Seol, Maury A. Nussbaum, Michael L. Madigan*
Department of Industrial and Systems Engineering, Virginia Tech, 250
Durham Hall (0118), Blacksburg, VA 24061, USA
* Department of Engineering Science and Mechanics, Virginia Tech,
Mail Code 0219, Blacksburg, VA 24061, USA

Measures of human sway during upright standing are frequently used as indirect measures of the postural control system. The objective of this study was to assess the reliability of several center-of-pressure-based sway measures. Test-retest reliability was determined both within-days (under different visual and surface conditions) and between-days (5 different days). Intraclass Correlation Coefficient (ICC), Standard Error of Measurement (SEM), Minimal Metrically Detectable Change (MMDC), and Coefficient of Variation (CV) were used as reliability indices. Mean velocity was found to be the most reliable measure among those evaluated. Overall, between-day reliability indices were better than within-day reliability. When subjects were standing on a compliant surface, several measures exhibited better reliability under 'eyes open' condition. These results can be useful in guiding the selection of postural sway measures and appropriate number of replications to improve sway assessment.

INTRODUCTION

Occupational falls are a major cause of work-related injuries and fatalities in contemporary industry. Decrements in balance control have been considered as an important factor in the probability of a fall (Redfern et al., 1997). Performance of the postural controls system is commonly assessed in the laboratory using a variety of sway measures. These sway measures are determined either from ground reaction forces or directly by evaluating body movements. The most commonly used measures are various parameters of the center of pressure (COP) obtained during quiet stance, as they are a simple and efficient way of examining multisensory integration in postural stability. Previous studies have shown that traditional COP-based measures (ellipse area, mean COP velocity, etc.) ignore the dynamic characteristics of COP movement and give few insights into postural control (Collins et al., 1995). To overcome this limitation, some advanced approaches have been described, including Hurst rescaled range analysis ($H_{R/S}$) and detrended fluctuation analysis (DFA). Some of these modern methods have been reported as more sensitive to effects on postural stability (e.g. fractal COP measures are more sensitive to the age-related changes; Norris et al., 2005). In general, selection of a sway measure should be motivated by several factors, including the sensitivity and reliability of the measure, the latter of which is the focus here.

Although COP-based methods are widely used in studies on postural control, the reliability of these measures is still not fully explored. Several studies have determined the reliability of the traditional COP-based measures (Benvenuti et al., 1999; Corriveau et al., 2000; Baratto et al., 2002). However, no study has reported the reliability of Hurst exponent and DFA that have been designed to reveal the hidden fractal properties of time series. Existing reliability studies have obtained sway measures during quiet stance for various length of time (40 sec ~ 120 sec), and for different test conditions (e.g. vision) and subject characteristics (e.g. age or disease). COP mean velocity has usually been reported as the most reliable measure among traditional measures, and within-day results appears to have better reliability than between-day results. In the present study, we assessed within- and between-day reliability of both traditional COP measures as well as several advanced measures in healthy young subjects using ICC, SEM, MMDC, and CV. Comprehensive assessment of reliability in this manner was expected to be improved over earlier studies in which different aspects of reliability have been determined separately. Results from this were intended to facilitate better selection of COP-based sway measures in future studies.

METHODS

Participants

Sixteen healthy individuals (8 males and 8 females) participated in the study. Their mean (SD) age was 21.5 (2.5), and they had no self-reported injuries, illnesses, musculoskeletal disorders, or occurrences of falls in the past year. All completed an informed consent procedure approved by the Virginia Tech Institutional Review Board.

Experimental Procedures

Day 1: Within-day reliability. Short-term test-retest reliability of sway measures was determined under four different conditions, while participants stood barefoot on a force platform (AMTI OR6-7-1000, Watertown, Massachusetts, USA) as still as possible, with arms at their side. During sway measurements participants' eyes were closed (EC) or opened (EO). During EO, they were asked to stare at a cross-mark that was 75 cm away from their eyes. The standing surface was either hard (HS) or soft (SS). Soft foam covered the force platform in SS, and the height of the soft surface was 23 mm. Three replications of each of the four conditions (2 vision x 2 surface) were performed, and the order of the 12 total trials was fully randomized, with at least one minute between each. Sway trials were conducted in a closed room to eliminate any noise or other disturbances. Repeatability of foot placement between trials was maintained by outlining the feet on poster board placed on top of the force plate. Each trial lasted 75 seconds, with the initial 10 s and last 5 s removed. Triaxial ground reaction forces and moments were sampled at 100 Hz and subsequently low-pass filtered (Butterworth, 5 Hz cut-off frequency, 2nd order, zero lag) and transformed to obtain COP values (Winter, 1995).

Day 1 ~ Day 5: Between-day reliability. Longer-term reliability was assessed by repeating a portion of the procedures described above 5 times (Day 1 ~ Day 5). This was done in the same laboratory environment, by the same evaluators, under the same conditions, and with a between-day interval of 2-14 days according to the availability of the participants. Between-days reliability was assessed only for the EC/HS condition. Day 1 data were obtained from the earlier within-day reliability study, while data for Days 2-5 were obtained from three replications of the EC/HS condition. Reliability was determined for the mean of the three replications obtained in each Day.

Quantifying reliability

Several methods were used to quantify reliability here, in order to obtain a more comprehensive evaluation. ICC was chosen to assess the proximity of the data points to a straight line which has a gradient of 1 and passes the origin of a coordinate system. SEM was selected to test absolute reliability. This can be used to distinguish low retest reliability coefficient caused by variability within subjects from low coefficients caused by a narrow range of values within the study sample (Henriksen et al., 2004). The CV was calculated because it is useful as a descriptive tool and is independent of the units of measurement.

ICCs were calculated using variance components from a two-way ANOVA model. ICC ranges from 0 (no reliability) to 1 (perfect reliability) and was interpreted using the following criteria: 0.00-0.39 poor, 0.40-0.59 fair, 0.60-0.74 good and 0.75-1.00 excellent (Cincchetti & Soarrow, 1981). ICC (2,1) was used for both within- and between-days reliability (Shrout & Fleiss, 1979). SEM (Stratford & Goldsmith, 1977) was calculated as SD $\sqrt{(1-ICC)}$, where SD is the standard deviation. A high SEM indicates a high level of error and implies non-reproducibility of tested values. MMDC, defined as the 95% confidence interval of the standard error of the measurement (±1.96 SEM) was used to distinguish the true value from noise caused by the inherent variability of the measurement (Lafond et al., 2004). CV (Samson & Crowe, 1996) was used to detect random measurement error and to compare different variables by scaling the SEM to the mean, and was calculated as [(SEM/mean)x100].

The dependent variables used in the study were COP-based sway measures derived from the stabilograms, which are the COP trajectories collected during sway trials. The following COP based measures were determined: 1) Ellipse area (Prieto et al., 1996); 2) Sway area; 3) Mean power frequency (Hasan et al., 1996); 4) Mean velocity; 5) RMS distance; 6) Hurst exponent (Duarte & Zatsiorsky., 2000); and 7) DFA scaling exponent (Norris et al., 2005).

RESULTS

The ICC, SEM, MMDC and CV results are shown in Table 1 and Table 2. Ellipse area appeared to have good within-day reliability and excellent between-day reliability. SEM and CV values of ellipse area were relatively high. Sway area had fair-good within-day reliability and excellent between-day reliability, according to the ICC values. The CV values were relatively high. For ellipse area and sway area, between-day reliability indices were better than within-day. The mean power frequency (AP and ML) showed poor-good within-day reliability and good between-day reliability. ICC values for mean velocity (AP and ML) were quite high and were accompanied by low SEM and CV values. Again, between-day indices showed better (excellent) reliability than within-day indices (good-excellent). RMS distance showed fair-good within-day reliability and excellent between-day reliability. SEM values for the RMS distance were higher in EC condition when compared to EO condition. Hurst exponents had poor-fair within-day reliability and fair between-day reliability. DFA scaling exponents had poor-good within-day reliability and good between-day reliability. CV values were very low for Hurst exponents and DFA scaling exponents. Under the hard surface condition, most dependent measures (except sway area) showed better reliability for eyes closed than eyes opened. Under the soft surface condition, ellipse area, sway area, mean velocity AP, and Hurst exponent AP showed better reliability when subjects' eyes were opened than closed.

Table 1 Within-day and between-day ICCs, SEMs, MMDC, and CV of traditional COP measures

		Within-day				Between-day
		EC, HS	EO, HS	EC, SS	EO, SS	EC,HS
Ellipse area	Mean (mm^2)	686	308	679	344	734
	ICC (2,1)	0.65	0.61	0.68	0.73	0.78
	SEM (mm^2)	262.4	103.8	208.1	116.4	230.0
	CV (%)	38.2	33.7	30.6	33.8	31.3
	MMDC (mm^2)	±514.3	±203.5	±407.8	±228.2	±450.8
Sway area	Mean (mm^2/s)	41.2	16.3	45.4	18.6	41.6
	ICC (2,1)	0.56	0.67	0.59	0.67	0.81
	SEM (mm^2/s)	13.3	4.3	14.5	5.6	8.9
	CV (%)	32.2	26.5	31.8	30.0	21.5
	MMDC (mm^2/s)	±26.0	±8.5	±28.3	±10.9	±17.5
Mean power frequency (AP)	Mean (Hz)	0.20	0.11	0.23	0.12	0.19
	ICC (2,1)	0.58	0.01	0.60	0.30	0.65
	SEM (Hz)	0.06	0.05	0.05	0.04	0.04
	CV (%)	29.1	41.6	20.5	34.7	21.8
	MMDC (Hz)	±0.12	±0.09	±0.09	±0.08	±0.08
Mean power frequency (ML)	Mean (Hz)	0.27	0.24	0.31	0.27	0.25
	ICC (2,1)	0.68	0.50	0.31	0.06	0.62
	SEM (Hz)	0.06	0.07	0.09	0.08	0.06
	CV (%)	21.6	29.8	27.7	31.2	24.3
	MMDC (Hz)	±0.11	±0.14	±0.17	±0.16	±0.12
Mean velocity (AP)	Mean (mm/s)	9.0	5.3	10.1	5.4	8.8
	ICC (2,1)	0.84	0.67	0.64	0.83	0.86
	SEM (mm/s)	0.9	0.8	1.4	0.6	0.9
	CV (%)	9.8	14.3	13.4	10.5	10.5
	MMDC (mm/s)	±1.7	±1.5	±2.7	±1.1	±1.8
Mean velocity (ML)	Mean (mm/s)	11.7	5.9	12.9	6.3	11.5
	ICC (2,1)	0.83	0.78	0.70	0.67	0.84
	SEM (mm/s)	1.3	0.7	2.0	0.8	1.7
	CV (%)	11.3	12.0	15.4	13.3	15.1
	MMDC (mm/s)	±2.6	±1.4	±3.9	±1.7	±3.4
RMS distance	Mean (mm)	8.6	6.2	8.6	6.6	8.8
	ICC (2,1)	0.55	0.46	0.64	0.49	0.79
	SEM (mm)	1.8	1.2	1.5	1.4	1.2
	CV (%)	20.9	19.9	17.9	21.4	13.5
	MMDC (mm)	±3.5	±2.4	±3.0	±2.8	±2.3

Table 2 Within-day and between-day ICCs, SEMs, MMDC, and CV of fractal-based COP measures

		Within-day				Between-day
		EC, HS	EO, HS	EC, SS	EO, SS	EC,HS
Hurst exponent (AP)	Mean	0.93	0.96	0.92	0.96	0.93
	ICC (2,1)	0.56	0.16	0.41	0.43	0.58
	SEM	0.01	0.01	0.02	0.01	0.01
	CV (%)	1.5	1.5	1.7	1.2	1.2
	MMDC	±0.03	±0.03	±0.03	±0.02	±0.02
Hurst exponent (ML)	Mean	0.91	0.92	0.90	0.91	0.92
	ICC (2,1)	0.60	0.41	0.43	0.01	0.58
	SEM	0.02	0.02	0.02	0.02	0.02
	CV (%)	2.0	1.8	2.5	2.6	1.8
	MMDC	±0.04	±0.03	±0.04	±0.05	±0.03
DFA (AP)	Mean	1.37	1.49	1.36	1.50	1.40
	ICC (2,1)	0.65	0.38	0.69	0.50	0.69
	SEM	0.04	0.06	0.04	0.04	0.04
	CV (%)	3.4	3.8	2.6	2.7	2.9
	MMDC	±0.09	±1.1	±0.07	±0.08	±0.08
DFA (ML)	Mean	1.34	1.34	1.31	1.32	1.36
	ICC (2,1)	0.70	0.62	0.53	0.20	0.67
	SEM	0.05	0.08	0.06	0.07	0.05
	CV (%)	3.7	5.8	4.3	5.1	3.8
	MMDC	±0.10	±0.15	±0.11	±0.13	±0.10

DISCUSSION

The main objective of this study was to assess within- and between-day test-retest reliability of several postural sway measures. Overall, the results showed that COP-based measures have different levels of reliability. Among the measures obtained, mean velocity was the most reliable COP-based measure, in accordance with previous studies (Lafond et al., 2004; Cornilleau-Pérès et al., 2005; Raymakers, Samson, & Verhaar, 2005). Lafond et al. (2004) suggested the high reliability of mean velocity makes this the most discriminating measure for age-related and visual changes.

Although it has been suggested that Hurst and DFA scaling exponents had some advantages compared with the traditional COP measures (Norris, et al., 2005), no study has reported the reliability of Hurst exponent and DFA scaling exponent. In our study, the ICC values of these measures showed poor-good reliability but the low SEM and CV values demonstrated quite good reliability. The relatively low ICC values may be explained by the little variation of Hurst and DFA scaling exponents. ICC is determined as the ratio of the variance between subjects to the total variance (Bruton, et al., 2000). The little variation of Hurst and DFA measures resulted in the small numerator of ICC calculation and leads to the relatively low ICC values.

Ellipse area, sway area, mean velocity AP, and the Hurst exponent AP showed better reliability when participants' eyes were opened than closed under the soft surface condition. Disabled visual function could have contributed to decreased postural stability and larger variation in COP-based measures. This effect might have been magnified in the soft surface condition (Redfern, Moor, & Yarsky, 1997; Simeonov & Hsiao, 2001). Most of dependent measures used in our study were more variable for eyes opened than eyes closed under the hard surface condition. This interesting result may be attributed to the randomization of four standing conditions. Lafond (2006) indicated that the reliability of COP measures should be assessed by testing subjects on successive trials of the same condition or in the same order.

Overall, between-day reliability indices were better than within-day reliability. The highest (excellent) ICC values were obtained for mean velocity AP and ML (0.86 and 0.84 respectively) followed by sway area (0.81), RMS distance (0.79), and ellipse area (0.78). A possible cause is that we used the mean of three values to calculate between-day reliability and that data for within-day assessment was collected on the first day of the experiment.

In this study, reliability of fractal-based COP measures has been determined, in addition to traditional

COP measures. The fractal-based COP measures did not show better reliability than the traditional ones. The results of this study can contribute toward choosing reliable postural sway measures and in deciding reasonable methods of postural sway testing such as duration, replications, and environmental conditions.

REFERENCES

Baratto, L., Morasso, P.G., Re, C., & Spida, G. (2002). A new look at posturographic analysis in the clinical context: sway-density versus other parameterization techniques. Motor Control, 6, 246-270.

Benvenuti, F., Mecacci, R., Gineprari, I., Bandinelli, S., Benvenuti, E., Ferrucci, L., Baroni, A., Rabufetti, M., Hallett, M., Dambrosia, J.M., & Stanhope, S.J. (1999). Kinematic characteristics of standing disequilibrium: reliability and validity of posturographic protocol. Archives of Physical Medicine and Rehabilitation, 80, 278-287.

Bruton, Al, Conway, J.H., & Holgate, S.T. (2000) Reliability: What is it, and how is it measured? Physiotherapy, 86, 94-99.

Cincchetti, D.V., & Sparrow, S.S. (1981). Developing criteria for establishing Interrater reliability of specific items: application to assessment of adaptive behavior. American Journal of Mental Deficiency, 86, 127-137.

Collins, J.J., De Luca, C.J., Burrows, A., & Lipsitz, L.A. (1995). Age-related changes in open-loop and closed-loop postural control mechanisms. Experimental Brain Research, 104, 480-492.

Cornilleau-Pérès, V., Shabana, N., Droulez, J., Goh, J.C.H., Lee, G.S.M., & Chew, P.T.K. (2005). Measurement of the visual contribution to postural steadiness from the COP movement: methodology and reliability. Gait & Posture, 22, 96-106.

Corriveau, H., Hébert, R., Prince, F., & Raiche, M. (2000). Intrasession reliability of the "center of pressure minus center of mass" variable of postural control in the healthy elderly. Archives of Physical Medicine and Rehabilitation, 81, 45-48.

Duarte, M., & Zatsiorsky, V.M. (2000). On the fractal properties of natural human standing. Neuroscience Letters, 283, 173-176.

Hasan, S.S., Robin, D.W., Szurkus, D.C., Ashmead, D.H., Peterson, S.W., & Shiavi, R.G. (1996). Simultaneous measurement of body center of pressure and center of gravity during upright stance. Part II: Amplitude and frequency data. Gait & Posture, 4, 11-20.

Henriksen, M., Lund, H., Moe-Nilssen, R., Bliddal, H., & Danneskiod-Samsoe, B. (2004). Test-retest reliability of trunk accelerometric gait analysis. Gait Posture,19, 288-297.

Lafond, D., Corriveau, H., Hébert, R., Prince, F., & Raiche, M. (2004). Intrasession reliability of the center of pressure measures of postural steadiness in healthy elderly people. Archives of Physical Medicine and Rehabilitation, 85, 896-901.

Lafond, D. (2006) Reliability of center of pressure measures of postural steadiness [letter]. Archives of Physical Medicine and Rehabilitation, 87, 308.

Norris, J.A., Marsh, A.P., Smith, I.J., Kohut, R.I., & Miller M.E. (2005). Ability of static and statistical mechanics posturographic measures to distinguish between age and fall risk. Journal of Biomechanics, 38, 1263-1272.

Prieto, T.E., Myklebust, J.B., Hoffmann, R.G., Lovett, E.G., & Myklebust, B.M. (1996). Measures of postural steadiness: differences between healthy young and elderly adults. IEEE Transactions on Biomedical Engineering, 43, 956-966.

Raymakers, J.A., Samson, M.M., Verhaar, H.J.J. (2005). The assessment of body sway and the choice of stability parameter(s). Gait and Posture, 21, 48-58.

Redfern, M.S., Moore, P.L., & Yarsky, C.M. (1997). The influence of flooring on standing balance among older persons. American Journal of Hygiene Association Journal, 61, 700-708.

Samson, M., Crowe, A. (1996). Intra-subject inconsistencies in quantitative assessment of body sway. Gait and Posture, 4, 252-257.

Shrout, P.E., Fleiss, J.L. (1979). Intraclass correlation: uses in assessing rater reliability. Psycological Bulletin, 86, 420-428.

Simeonov, P., Hsiao, H. (2001). Height, surface firmness, and visual reference effects on balance control. Injury Prevention, 7 (Suppl 1), i50-53.

Stratford, P.W., Goldsmith, C.H. (1997). Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. Physical Therapy, 77, 745-750.

Winter, D.A. (1995). Human balance and posture control during standing and walking. Gait and Posture, 3, 193-214.